OPEN ACCESS Freely available online
PLoS GENETICS

# A Genome-Wide Association Meta-Analysis of Circulating Sex Hormone–Binding Globulin Reveals Multiple Loci Implicated in Sex Steroid Hormone Regulation

Andrea D. Coviello[1,2,3☯], Robin Haring[4☯], Melissa Wellons[5☯], Dhananjay Vaidya[6☯], Terho Lehtimäki[7☯], Sarah Keildson[8☯], Kathryn L. Lunetta[9], Chunyan He[10,11], Myriam Fornage[12], Vasiliki Lagou[8,13], Massimo Mangino[14], N. Charlotte Onland-Moret[15], Brian Chen[16], Joel Eriksson[17], Melissa Garcia[18], Yong Mei Liu[19,20], Annemarie Koster[21], Kurt Lohman[19], Leo-Pekka Lyytikäinen[7], Ann-Kristin Petersen[22], Jennifer Prescott[23,24], Lisette Stolk[25,26], Liesbeth Vandenput[17], Andrew R. Wood[27], Wei Vivian Zhuang[9], Aimo Ruokonen[28], Anna-Liisa Hartikainen[29], Anneli Pouta[30], Stefania Bandinelli[31], Reiner Biffar[32], Georg Brabant[33], David G. Cox[34,35], Yuhui Chen[8], Steven Cummings[36], Luigi Ferrucci[37], Marc J. Gunter[35], Susan E. Hankinson[24,38,39], Hannu Martikainen[29], Albert Hofman[26,40], Georg Homuth[41], Thomas Illig[42,43], John-Olov Jansson[17], Andrew D. Johnson[3], David Karasik[44], Magnus Karlsson[45], Johannes Kettunen[46,47], Douglas P. Kiel[44], Peter Kraft[48], Jingmin Liu[49], Östen Ljunggren[50], Mattias Lorentzon[17], Marcello Maggio[51], Marcello R. P. Markus[52], Dan Mellström[17], Iva Miljkovic[53], Daniel Mirel[54], Sarah Nelson[55], Laure Morin Papunen[29], Petra H. M. Peeters[15], Inga Prokopenko[8,13], Leslie Raffel[56], Martin Reincke[57], Alex P. Reiner[58], Kathryn Rexrode[59], Fernando Rivadeneira[25,26], Stephen M. Schwartz[60], David Siscovick[60], Nicole Soranzo[14,61], Doris Stöckl[62,63], Shelley Tworoger[24,39], André G. Uitterlinden[25,26,40], Carla H. van Gils[15], Ramachandran S. Vasan[1,3], H.-Erich Wichmann[64,65,66], Guangju Zhai[14,67], Shalender Bhasin[2], Martin Bidlingmaier[57], Stephen J. Chanock[68], Immaculata De Vivo[23,24], Tamara B. Harris[21], David J. Hunter[23,24], Mika Kähönen[69], Simin Liu[70], Pamela Ouyang[71], Tim D. Spector[14], Yvonne T. van der Schouw[15], Jorma Viikari[72], Henri Wallaschofski[4], Mark I. McCarthy[8,73,74], Timothy M. Frayling[27], Anna Murray[27], Steve Franks[75], Marjo-Riitta Järvelin[76,77,78,79¶], Frank H. de Jong[25¶], Olli Raitakari[80¶], Alexander Teumer[41¶], Claes Ohlsson[17¶], Joanne M. Murabito[3,81¶*], John R. B. Perry[8,14,27¶*]

**1** Section of Preventive Medicine and Epidemiology, Boston University School of Medicine, Boston, Massachusetts, United States of America, **2** Section of Endocrinology, Diabetes, and Nutrition, Boston University School of Medicine, Boston, Massachusetts, United States of America, **3** National Heart, Lung, and Blood Institute's The Framingham Heart Study, Framingham, Massachusetts, United States of America, **4** Institute for Clinical Chemistry and Laboratory Medicine, University Medicine, Ernst-Moritz-Arndt University of Greifswald, Greifswald, Germany, **5** Department of Medicine and Department of Obstetrics and Gynecology, The University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **6** Department of Medicine, Johns Hopkins University, Baltimore, Maryland, United States of America, **7** Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland, **8** Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, **9** Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States of America, **10** Department of Public Health, Indiana University School of Medicine, Indianapolis, Indiana, United States of America, **11** Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, Indiana, United States of America, **12** University of Texas Health Sciences Center at Houston, Houston, Texas, United States of America, **13** Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, Oxford, United Kingdom, **14** Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom, **15** Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, **16** Program on Genomics and Nutrition and the Center for Metabolic Disease Prevention, School of Public Health, University of California Los Angeles, Los Angeles, California, United States of America, **17** Center for Bone and Arthritis Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, **18** Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, Maryland, United States of America, **19** Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States of America, **20** Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, North Carolina, United States of America, **21** Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, Maryland, United States of America, **22** Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany, **23** Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America, **24** Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts, United States of America, **25** Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands, **26** Netherlands Consortium of Healthy Aging, Rotterdam, The Netherlands, **27** Genetics of Complex Traits, Peninsula Medical School, University of Exeter, Exeter, United Kingdom, **28** Institute of Diagnostics, University of Oulu, Oulu, Finland, **29** Department of Obstetrics and Gynecology, University Hospital of Oulu, Oulu, Finland, **30** National Institute for Health and Welfare and Institute of Health Sciences, University of Oulu, Oulu, Finland, **31** Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy, **32** Department of Prosthetic Dentistry, Gerostomatology, and Dental Materials, University of Greifswald, Greifswald, Germany, **33** Experimental and Clinical Endocrinology, University of Lübeck, Lübeck, Germany, **34** Cancer Research Center of Lyon, INSERM U1052, Lyon, France, **35** Department of Epidemiology and Biostatistics, School of Public Health, Imperial College, London, United Kingdom, **36** California Pacific Medical Center, San Francisco, California, United States of America, **37** Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, Baltimore, Maryland, United States of America, **38** Division of Biostatistics and Epidemiology, University of Massachusetts, Amherst, Massachusetts, United States of America, **39** Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America, **40** Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands, **41** Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany, **42** Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany, **43** Hannover Unified Biobank, Hannover Medical School, Hannover, Germany, **44** Hebrew SeniorLife Institute for Aging Research and Harvard Medical School, Boston, Massachusetts, United States of America, **45** Clinical and Molecular Osteoporosis Research Unit, Department of Clinical Sciences and Department of Orthopaedics, Lund University, Malmö, Sweden, **46** Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland, **47** Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki,

Finland, **48** Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America, **49** Women's Health Initiative Clinical Coordinating Center, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **50** Department of Medical Sciences, University of Uppsala, Uppsala, Sweden, **51** Department of Internal Medicine and Biomedical Sciences, Section of Geriatrics, University of Parma, Parma, Italy, **52** Institute for Community Medicine, University of Greifswald, Greifswald, Germany, **53** University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **54** Gene Environment Initiative, Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, Massachusetts, United States of America, **55** Department of Biostatistics, University of Washington, Seattle, Washington, United States of America, **56** Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, **57** Medizinische Klinik and Poliklinik IV, Ludwig-Maximilians University, Munich, Germany, **58** Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **59** Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **60** Cardiovascular Health Research Unit, Department of Epidemiology, University of Washington, Seattle, Washington, United States of America, **61** Human Genetics, Wellcome Trust Sanger Institute, Hinxton, United Kingdom, **62** Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany, **63** Department of Obstetrics and Gynaecology, Ludwig-Maximilians-University, Munich, Germany, **64** Institute of Epidemiology I, Helmholtz Zentrum München, Neuherberg, Germany, **65** Institute of Medical Informatics, Biometry, and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany, **66** Klinikum Großhadern, Munich, Germany, **67** Discipline of Genetics, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada, **68** Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, United States of America, **69** Department of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland, **70** Program on Genomics and Nutrition, Department of Epidemiology, University of California Los Angeles, Los Angeles, California, United States of America, **71** Division of Cardiology, Johns Hopkins Bayview Medical Center, Baltimore, Maryland, United States of America, **72** Department of Medicine, Turku University Hospital and University of Turku, Turku, Finland, **73** Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, Churchill Hospital, Oxford, United Kingdom, **74** Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, United Kingdom, **75** Institute of Reproductive and Developmental Biology, Imperial College London, London, United Kingdom, **76** Department of Biostatistics and Epidemiology, School of Public Health, MRC-HPA Centre for Environment and Health, Faculty of Medicine, Imperial College London, London, United Kingdom, **77** Institute of Health Sciences, University of Oulu, Oulu, Finland, **78** Biocenter Oulu, University of Oulu, Oulu, Finland, **79** National Institute of Health and Welfare, University of Oulu, Oulu, Finland, **80** Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital and Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland, **81** Section of General Internal Medicine, Boston University School of Medicine, Boston, Massachusetts, United States of America

## Abstract

Sex hormone-binding globulin (SHBG) is a glycoprotein responsible for the transport and biologic availability of sex steroid hormones, primarily testosterone and estradiol. SHBG has been associated with chronic diseases including type 2 diabetes (T2D) and with hormone-sensitive cancers such as breast and prostate cancer. We performed a genome-wide association study (GWAS) meta-analysis of 21,791 individuals from 10 epidemiologic studies and validated these findings in 7,046 individuals in an additional six studies. We identified twelve genomic regions (SNPs) associated with circulating SHBG concentrations. Loci near the identified SNPs included *SHBG* (rs12150660, 17p13.1, $p = 1.8\times10^{-106}$), *PRMT6* (*rs17496332*, *1p13.3*, $p = 1.4\times10^{-11}$), *GCKR* (*rs780093*, *2p23.3*, $p = 2.2\times10^{-16}$), *ZBTB10* (*rs440837*, *8q21.13*, $p = 3.4\times10^{-09}$), *JMJD1C* (*rs7910927*, *10q21.3*, $p = 6.1\times10^{-35}$), *SLCO1B1* (*rs4149056*, *12p12.1*, $p = 1.9\times10^{-08}$), *NR2F2* (*rs8023580*, *15q26.2*, $p = 8.3\times10^{-12}$), *ZNF652* (*rs2411984*, *17q21.32*, $p = 3.5\times10^{-14}$), *TDGF3* (*rs1573036*, *Xq22.3*, $p = 4.1\times10^{-14}$), *LHCGR* (*rs10454142*, *2p16.3*, $p = 1.3\times10^{-07}$), *BAIAP2L1* (*rs3779195*, *7q21.3*, $p = 2.7\times10^{-08}$), and *UGT2B15* (*rs293428*, *4q13.2*, $p = 5.5\times10^{-06}$). These genes encompass multiple biologic pathways, including hepatic function, lipid metabolism, carbohydrate metabolism and T2D, androgen and estrogen receptor function, epigenetic effects, and the biology of sex steroid hormone-responsive cancers including breast and prostate cancer. We found evidence of sex-differentiated genetic influences on SHBG. In a sex-specific GWAS, the loci 4q13.2-*UGT2B15* was significant in men only (men $p = 2.5\times10^{-08}$, women $p = 0.66$, heterogeneity $p = 0.003$). Additionally, three loci showed strong sex-differentiated effects: 17p13.1-*SHBG* and Xq22.3-*TDGF3* were stronger in men, whereas 8q21.12-*ZBTB10* was stronger in women. Conditional analyses identified additional signals at the *SHBG* gene that together almost double the proportion of variance explained at the locus. Using an independent study of 1,129 individuals, all SNPs identified in the overall or sex-differentiated or conditional analyses explained ~15.6% and ~8.4% of the genetic variation of SHBG concentrations in men and women, respectively. The evidence for sex-differentiated effects and allelic heterogeneity highlight the importance of considering these features when estimating complex trait variance.

**Citation:** Coviello AD, Haring R, Wellons M, Vaidya D, Lehtimäki T, et al. (2012) A Genome-Wide Association Meta-Analysis of Circulating Sex Hormone–Binding Globulin Reveals Multiple Loci Implicated in Sex Steroid Hormone Regulation. PLoS Genet 8(7): e1002805. doi:10.1371/journal.pgen.1002805

**Editor:** Greg Gibson, Georgia Institute of Technology, United States of America

**Received** February 20, 2012; **Accepted** May 19, 2012; **Published** July 19, 2012



**Funding:** This meta-analysis is a collaborative effort involving data from many individual studies and many sources of funding. The details of funding sources for each study are detailed in Text S1 as well as below. Framingham Heart Study (FHS): The phenotype-genotype association analyses were funded through grants from the NIA R21AG032598 (JM Murabito, KL Lunetta), R01HL094755 (AD Coviello, RS Vasan, S Bandinelli), and R01AG31206 (RS Vasan, S Bandinelli), R01 AR/AG 41398 (DP Kiel). This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II), funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. Gothenburg Osteoporosis and Obesity Determinants (GOOD) Study: Financial support was received from the Swedish Research Council (K2010-54X-09894-19-3, 2006-3832, and K2010-52X-20229-05-3), the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk foundation, and the European Commission grant HEALTH-F2-2008-201865-GEFOS. We would like to acknowledge Maria Nethander at the genomics core facility at University of Gothenburg for statistical analyses. We would also like to thank Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and #01 IG 07015 G for access to their grid resources. We would also like to thank Karol Estrada, Department ofInternal Medicine, Erasmus MC, Rotterdam, The Netherlands, for advice regarding

the grid resources. Health, Aging, and Body Composition (Health ABC) Study: This Health ABC Study was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was also supported in part by the Intramural Research Program of the National Institute on Aging, NIH, Bethesda, Maryland. Invecchiare in Chianti (InCHIANTI): The InCHIANTI study baseline (1998–2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336); the InCHIANTI Follow-up 1 (2001–2003) was funded by the U.S. National Institute on Aging (Contracts: N.1-AG-1-1 and N.1-AG-1-2111); the InCHIANTI Follow-ups 2 and 3 studies (2004–2010) were financed by the U.S. National Institute on Aging (Contract: N01-AG-5-0002), supported in part by the Intramural research program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland. JRB Perry is a Sir Henry Wellcome Postdoctoral Research Fellow (092447/Z/10/Z). Cooperative Health Research in the Region of Augsburg (KORA): The KORA research platform was initiated and financed by the Helmholtz Zentrum Munich, German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2 and NGFNPlus: 01GS0823). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ. Multi-Ethnic Study of Atherosclerosis (MESA): MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support is provided by grants and contracts N01 HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, and RR-024156. Funding support for the sex hormone dataset was provided by grants HL074406 and HL074338. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. Genotyping was performed at the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) and at Affymetrix (Santa Clara, California, USA) using the Affymetric Genome-Wide Human SNP Array 6.0. Northern Finland Birth Cohort 1966 Study (NFBC-66): NFBC1966 received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, Center of Excellence in Complex Disease Genetics, and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), the European Commission (EURO-BLCS, Framework 5 award QLG1-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, the Medical Research Council UK (G0500539, G0600705, PrevMetSyn/SALVE), and the Wellcome Trust (project grant GR069224). We acknowledge the support of U.S. National Heart, Lung, and Blood Institute grant HL087679 through the STAMPEED program; grants MH083268, GM053275-14, and U54 RR020278 from the U.S. National Institutes of Health; grant DMS-0239427 from the National Science Foundation; the Medical Research Council of the UK, EURO-BLCS, QLG1-CT-2000-01643 and the European Community's Seventh Framework Programme (FP7/2007–2013); ENGAGE project and grant agreement HEALTH-F4-2007-201413. The authors would like to thank the Center of Excellence in Common Disease Genetics of the Academy of Finland and Nordic Center of Excellence in Disease Genetics, the Sydantautisaatio (Finnish Foundation of Heart Diseases), the Broad Genotyping Center, D. Mirel, H. Hobbs, J. DeYoung, P. Rantakallio, M. Koiranen, and M. Isohanni for advice and assistance. Rotterdam study (RS1): The generation and management of GWAS genotype data for the Rotterdam Study are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810, and funding from the European Commision (HEALTH-F2-2008-201865, GEFOS; HEALTH-F2-2008-35627, TREAT-OA). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam; Netherlands Organization for the Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture, and Science; the Ministry for Health, Welfare, and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. We thank Pascal Arp, Mila Jhamai, Dr. Michael Moorhouse, Marijn Verkerk, and Sander Bervoets for their help in creating the GWAS database. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We would like to thank Dr. Tobias A. Knoch, Karol Estrada, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 Gfor access to their grid resources. Study of Health in Pomerania (SHIP): SHIP is part of the Community Medicine Research Net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), and the Ministry of Cultural Affairs, as well as the Social Ministry of the Federal State of Mecklenburg, West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg, West Pomerania. The University of Greifswald is a member of the "Center of Knowledge Interchange" program of the Siemens AG. This work is also part of the research project Greifswald Approach to Individualized Medicine (GANI_MED). The GANI_MED consortium is funded by the Federal Ministry of Education and Research and the Ministry of Cultural Affairs of the Federal State of Mecklenburg, West Pomerania (03IS2061A). The SHBG reagents used were sponsored by Siemens Healthcare Diagnostics, Eschborn, formerly DPC Biermann GmbH, Bad Nauheim, Germany. Novo Nordisk provided partial grant support for the determination of serum samples and data analysis. R Haring received honorarium for lectures by Bayer Pharma AG. H Wallaschofski has received research grants from Novo Nordisk and Pfizer for research unrelated to the contents of this manuscript and honorarium for lectures by Bayer Pharma AG. TWINS UK: The study was funded by the Wellcome Trust, European Community's Seventh Framework Programme (FP7/2007–2013) grant agreement HEALTH-F2-2008-201865-GEFOS and (FP7/2007–2013), ENGAGE project grant agreement HEALTH-F4-2007-201413, and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's and St. Thomas' NHS Foundation Trust in partnership with King's College London. TD Spector is an NIHR senior Investigator. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project grant (G20234). The authors acknowledge the funding and support of the National Eye Institute via an NIH/CIDR genotyping project (PI: Terri Young). We thank the staff from the Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation, quality control, and genotyping led by Leena Peltonen and Panos Deloukas; Le Centre National de Génotypage, France, led by Mark Lathrop, for genotyping; Duke University, North Carolina, USA, led by David Goldstein, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki, led by Aarno Palotie. Genotyping was also performed by CIDR as part of an NEI/NIH project grant. The Cardiovascular Risk in Young Finns Study (YFS): The Young Finns Study has been financially supported by the Academy of Finland: grants 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere, and Turku University Hospital Medical Funds (grant 9M048 for 9N035 for TeLeht); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation; Tampere Tuberculosis Foundation; and Emil Aaltonen Foundation (T Lehtimäki). The expert technical assistance in the statistical analyses by Irina Lisinen and Ville Aalto are gratefully acknowledged. Women's Health Initiative (WHI): Genotyping was performed at the Broad Institute (Cambridge, MA) through the NHGRI-funded Genomics and Randomized Clinical Network (U01 HG005152) or GARNET. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts N01WH22110, 24152, 32100–2, 32105–6, 32108–9, 32111–13, 32115, 32118–32119, 32122, 42107–26, 42129–32, and 44221. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A listing of WHI investigators can be found at http://www.whiscience.org/publications/WHI_investigators_shortlist.pdf. Coronary Artery Risk Development in Young Adults (CARDIA) Women's Study: The CARDIA study is funded by contracts N01-HC-95095, N01-HC-48047, N01-HC-48048, N01-HC-48049, N01-HC-48050, N01-HC-45134, N01-HC-05187, N01-HC-45205, and N01-HC-45204 and by the CARDIA Women's study by R01-HL065611 from the National Heart, Lung, and Blood Institute to the CARDIA investigators. Genotyping of the CARDIA participants was supported by grants U01-HG-004729, U01-HG-004446, and U01-HG-004424 from the National Human Genome Research Institute. Statistical analyses were supported by grants U01-HG-004729 and R01-HL-084099 to M Fornage. M Wellons is supported by the Career Development Award 5-K23-HL087114. European Prospective Investigation into Cancer and Nutrition (Prospect-EPIC): The Prospect-EPIC study was funded by "Europe against Cancer" Programme of the European Commission (SANCO); the Dutch Ministry of Health, Welfare, and Sports (VWS); and ZONMw. Osteoporotic fractures in men (MrOS) study Sweden: Financial support was received from the Swedish Research Council (2006-3832), the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk Foundation, and the European Commission grant HEALTH-F2-2008-201865-GEFOS. Nurses' Health Study (NHS): The NHS breast cancer GWAS was performed as part of the Cancer Genetic Markers of Susceptibility (CGEMS) initiative of the NCI. We particularly acknowledge the contributions of R. Hoover, A. Hutchinson, K. Jacobs and G. Thomas. The current research is supported by CA87969, CA49449, CA40356, CA128034, and U01-CA98233 from the National Cancer Institute. We acknowledge the study participants in the NHS for their contribution in making this study possible. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* E-mail: John.perry@pms.ac.uk (JRB Perry); murabito@bu.edu (JM Murabito)

☯ These authors contributed equally to this work.

¶ These authors also contributed equally to this work.

## Author Summary

Sex hormone-binding globulin (SHBG) is the key protein responsible for binding and transporting the sex steroid hormones, testosterone and estradiol, in the circulatory system. SHBG regulates their bioavailability and therefore their effects in the body. SHBG has been linked to chronic diseases including type 2 diabetes and to hormone-sensitive cancers such as breast and prostate cancer. SHBG concentrations are approximately 50% heritable in family studies, suggesting SHBG concentrations are under significant genetic control; yet, little is known about the specific genes that influence SHBG. We conducted a large study of the association of SHBG concentrations with markers in the human genome in ~22,000 white men and women to determine which loci influence SHBG concentrations. Genes near the identified genomic markers in addition to the *SHBG* protein coding gene included *PRMT6*, *GCKR*, *ZBTB10*, *JMJD1C*, *SLCO1B1*, *NR2F2*, *ZNF652*, *TDGF3*, *LHCGR*, *BAIAP2L1*, and *UGT2B15*. These genes represent a wide range of biologic pathways that may relate to SHBG function and sex steroid hormone biology, including liver function, lipid metabolism, carbohydrate metabolism and type 2 diabetes, and the development and progression of sex steroid hormone-responsive cancers.


## Introduction

Sex hormone-binding globulin (SHBG) is a protein secreted mainly by the liver that binds to the sex steroids, testosterone, dihydrotestosterone, and estradiol, transports them in the circulation, and influences their action in target tissues by regulating their bioavailability. SHBG thereby influences the expression of sex hormone sensitive phenotypes including sexual characteristics and reproductive function in men and women [1]. In addition to regulating sex steroid hormone effects, SHBG may exert independent effects through its own receptor [2]. Variation in SHBG concentration has also been associated with various chronic diseases including cancers [3], polycystic ovary syndrome (PCOS) [4,5] and type 2 diabetes (T2D) [6,7]. Although SHBG is estimated to have a heritable component (~50%) [8], little is known about the genetic regulation of SHBG. Polymorphisms at the *SHBG* gene locus have been associated with SHBG concentrations [9,10], but much remains unknown about specific genetic variants that may determine circulating SHBG concentrations. Identifying genetic factors that influence SHBG may provide insights into the biology of sex steroid hormone regulation, metabolism and tissue effects that underlie their relationship with chronic diseases such as T2D as well as hormone-sensitive cancers such as breast and prostate cancer.

## Results

We identified nine loci associated with SHBG concentrations at the genome-wide significance threshold of $p = 5\times10^{-8}$ (Table 1 and Figure 1) in a genome-wide association study (GWAS) meta-analysis of circulating SHBG concentrations in 21,791 men and women from 10 studies (Table S1). All nine lead SNPs at these loci had effects in the same direction (seven with $p<0.05$) in the validation dataset of 7,046 men and women from six additional studies (Table S2). The strongest association was within the *SHBG* locus (rs12150660, $p = 2\times10^{-106}$). Together, these nine lead SNPs explained 7.2% of the genetic variance (assuming 50% heritability) in SHBG concentrations.

We next performed a series of additional analyses to explain more of the phenotypic variance (Figure 2). First, we hypothesized that genetic effects may be different in men and women, as SHBG concentrations are >50% higher in females than males, and may be differentially regulated between sexes. In a sex stratified analysis, three of the nine loci showed evidence of sex-differentiated effects at $p<0.02$ when we would not expect any signals to have reached this level of significance by chance. The associations at the 17p13.1-*SHBG* and Xq22.3 loci were stronger in males whereas the association at the 8q21.13 locus was stronger in females. To investigate the apparent differential sex effect for the X chromosome further we ran a recessive regression model for the X chromosome SNP rs1573036 in women in the Framingham Heart Study and found no association with SHBG suggesting the sex-differentiated effect is not the result of a recessive inheritance pattern. Sex stratified GWAS identified one novel signal in men, which showed no association in women (4q13.2: men $p = 2.5\times10^{-8}$, women $p = 0.66$, heterogeneity $p = 0.003$).

A series of conditional analyses were performed to identify statistically independent signals. At the *SHBG* locus, three apparently independent additional signals separate from the main index SNP were observed, based on low ($r^2<0.05$) pairwise correlations in HapMap (rs6258 $p = 2.7\times10^{-46}$, rs1625895 $p = 1.2\times10^{-14}$ and rs3853894 $p = 2.5\times10^{-11}$). A series of iterative conditional analyses (Table 2) involving SNPs at the *SHBG* locus generated a final regression model including six statistically independent *SHBG* SNPs. Four of these SNPs (#1–4 Table 2) retained GWS when conditioned against the other five, and two were nominally associated (SNP#5 $p = 0.0001$, SNP#6 $p = 0.01$). Re-running the GWAS meta-analysis adjusting for these six SNPs revealed evidence for three additional statistically independent (pairwise HapMap $r^2<0.01$) signals at the *SHBG* locus (SNP#7 $p = 1.5\times10^{-7}$, SNP#8 $p = 4.6\times10^{-5}$, SNP#9 $p = 9.9\times10^{-6}$) (Figure 3). There were also two additional *trans* signals located at 2p16.3 and 7q21.3 (Table 1). Although the 2p16.3 signal dropped below GWS when combined with follow-up samples ($p = 1\times10^{-7}$), the index SNP at 2p16.3 is ~300 kb away from a strong candidate gene, the luteinizing hormone receptor gene (*LHCGR*).

The majority of pair-wise correlations for the nine *SHBG* locus SNPs highlighted by our conditional analyses showed very low HapMap $r^2$ values. However, the pairwise D′ values are often high (Table S3) indicating that no or few recombination events have occurred between some SNPs, and that combinations of SNPs may be tagging un-typed variants on a common haplotype. To investigate this possibility, we performed more extensive analyses in a single study (NFBC1966, n = 4467). We used a denser set of SNPs imputed from the June 2011 version of the 1000 Genomes data and performed model selection analyses. Model selection identifies a set of SNPs that best explain phenotypic variation, while simultaneously penalizing each SNP included in this set, and therefore correlated SNPs tend to be excluded from the final model. These analyses consistently included at least seven SNPs in the model, although it is hard to estimate the false-negative rate of using the reduced sample size. While we are underpowered to accurately pinpoint the exact number of independent signals, these analyses support the results of the conditional analysis and suggest that multiple variants at the *SHBG* locus are independently associated with SHBG concentrations.

Data from an independent study, the InCHIANTI study, was used to calculate the proportion of genetic variance in SHBG concentrations explained when accounting for sex specific effects, the multiple signals of association at the *SHBG* locus, and the additional *trans* signals identified post conditional analysis. In men and women we explained ~15.6% and ~8.4% of the heritable

**Table 1.** SNPs representing loci associated with circulating SHBG concentrations.

| | | | | | | | | | Discovery Samples | | | Discovery+Follow-up | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Index SNP** | **Analysis** | **Region** | **Nr Gene** | **Chr** | **Position** | **Effect Allele** | **Other Allele** | **EAF** | **Beta** | **SE** | **p** | **Beta** | **SE** | **p** |
| rs17496332 | Main | 1p13.3 | PRMT6 | 1 | 107347898 | a | g | 0.67 | −0.026 | 0.0046 | 1.0E-08 | −0.028 | 0.0041 | 1.4E-11 |
| rs780093 | Main | 2p23.3 | GCKR | 2 | 27596107 | t | c | 0.40 | −0.033 | 0.0043 | 5.8E-14 | −0.032 | 0.0039 | 2.2E-16 |
| rs440837 | Main | 8q21.13 | ZBTB10 | 8 | 81624529 | a | g | 0.78 | −0.030 | 0.0052 | 6.7E-09 | −0.028 | 0.0047 | 3.4E-09 |
| rs7910927 | Main | 10q21.3 | JMJD1C | 10 | 64808916 | t | g | 0.51 | −0.044 | 0.0043 | 7.4E-25 | −0.048 | 0.0039 | 6.1E-35 |
| rs4149056 | Main | 12p12.1 | SLCO1B1 | 12 | 21222816 | t | c | 0.82 | 0.032 | 0.0057 | 1.5E-08 | 0.029 | 0.0052 | 1.9E-08 |
| rs8023580 | Main | 15q26.2 | NR2F2 | 15 | 94509295 | t | c | 0.72 | −0.029 | 0.0049 | 2.8E-09 | −0.03 | 0.0044 | 8.3E-12 |
| rs12150660 | Main | 17p13.1 | SHBG | 17 | 7462640 | t | g | 0.24 | 0.100 | 0.0053 | 1.2E-79 | 0.103 | 0.0047 | 1.8E-106 |
| rs2411984 | Main | 17q21.32 | ZNF652 | 17 | 44800750 | a | g | 0.28 | 0.034 | 0.0049 | 1.5E-12 | 0.033 | 0.0044 | 3.5E-14 |
| rs1573036 | Main | Xq22.3 | TDGF3 | 23 | 109706724 | t | c | 0.39 | 0.031 | 0.0043 | 5.1E-13 | 0.028 | 0.0037 | 4.1E-14 |
| rs10454142 | Conditional | 2p16.3 | LHCGR | 2 | 48499903 | t | c | 0.69 | 0.026 | 0.0047 | 2.8E-08 | 0.023 | 0.0044 | 1.3E-07 |
| rs3779195 | Conditional | 7q21.3 | BAIAP2L1 | 7 | 97831298 | a | t | 0.17 | −0.033 | 0.0057 | 1.2E-08 | −0.028 | 0.0051 | 2.7E-08 |
| rs293428 | Sex-specific | 4q13.2 | UGT2B15 | 4 | 69626371 | a | g | 0.69 | −0.023 | 0.0047 | 1.6E-06 | −0.019 | 0.0042 | 5.5E-06 |

| | Discovery Men | | | Discovery+Follow-up Men | | | Discovery Women | | | Discovery+Follow-up Women | | | Heterogeneity | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Index SNP** | **Beta** | **SE** | **p** | **Beta** | **SE** | **p** | **Beta** | **SE** | **p** | **Beta** | **SE** | **p** | **p** | **Sex** |
| rs17496332 | −0.027 | 0.0054 | 9.6E-07 | −0.027 | 0.0051 | 1.5E-07 | −0.02 | 0.008 | 0.003 | −0.029 | 0.0067 | 1.8E-05 | 0.79 | Females |
| rs780093 | −0.029 | 0.0052 | 1.8E-08 | −0.026 | 0.0049 | 7.0E-08 | −0.04 | 0.0076 | 1.3E-07 | −0.041 | 0.0063 | 8.6E-11 | 0.07 | Females |
| rs440837 | −0.021 | 0.0062 | 0.0009 | −0.019 | 0.0058 | 0.001 | −0.049 | 0.0093 | 1.2E-07 | −0.042 | 0.0078 | 7.2E-08 | 0.02 | Females |
| rs7910927 | −0.049 | 0.0051 | 5.3E-22 | −0.050 | 0.0048 | 1.2E-25 | −0.038 | 0.0075 | 6.4E-07 | −0.046 | 0.0063 | 1.7E-13 | 0.63 | Males |
| rs4149056 | 0.028 | 0.0067 | 0.00003 | 0.027 | 0.0063 | 1.5E-05 | 0.049 | 0.0103 | 0.000002 | 0.037 | 0.0086 | 1.7E-05 | 0.36 | Females |
| rs8023580 | −0.024 | 0.0057 | 0.00002 | −0.025 | 0.0054 | 5.1E-06 | −0.038 | 0.0087 | 0.00001 | −0.038 | 0.0071 | 7.8E-08 | 0.13 | Females |
| rs12150660 | 0.106 | 0.0063 | 1.8E-63 | 0.110 | 0.0058 | 3.7E-80 | 0.085 | 0.0094 | 1.8E-19 | 0.087 | 0.0077 | 5.8E-30 | 0.02 | Males |
| rs2411984 | 0.034 | 0.0058 | 5.9E-09 | 0.034 | 0.0054 | 2.3E-10 | 0.032 | 0.0084 | 0.0001 | 0.029 | 0.007 | 3.2E-05 | 0.54 | Males |
| rs1573036 | 0.04 | 0.0048 | 9.1E-17 | 0.035 | 0.0043 | 2.8E-16 | 0.012 | 0.0083 | 0.15 | 0.016 | 0.007 | 0.02 | 0.02 | Males |
| rs10454142 | | | | | | | | | | | | | | |
| rs3779195 | | | | | | | | | | | | | | |
| rs293428 | −0.032 | 0.0056 | 1.5E-08 | −0.029 | 0.0053 | 2.5E-08 | −0.005 | 0.0085 | 0.57 | −0.003 | 0.0071 | 0.66 | 0.003 | Males |

All SNPs are on the+strand and positions are based on build 36. EAF = 'effect allele frequency'. Beta units are per-allele effect estimates in natural log transformed nmol/L. Sex column gives the sex with the largest per-allele beta estimate. Missing values for conditional SNPs as sex-specific conditional analysis was not performed.

doi:10.1371/journal.pgen.1002805.t001

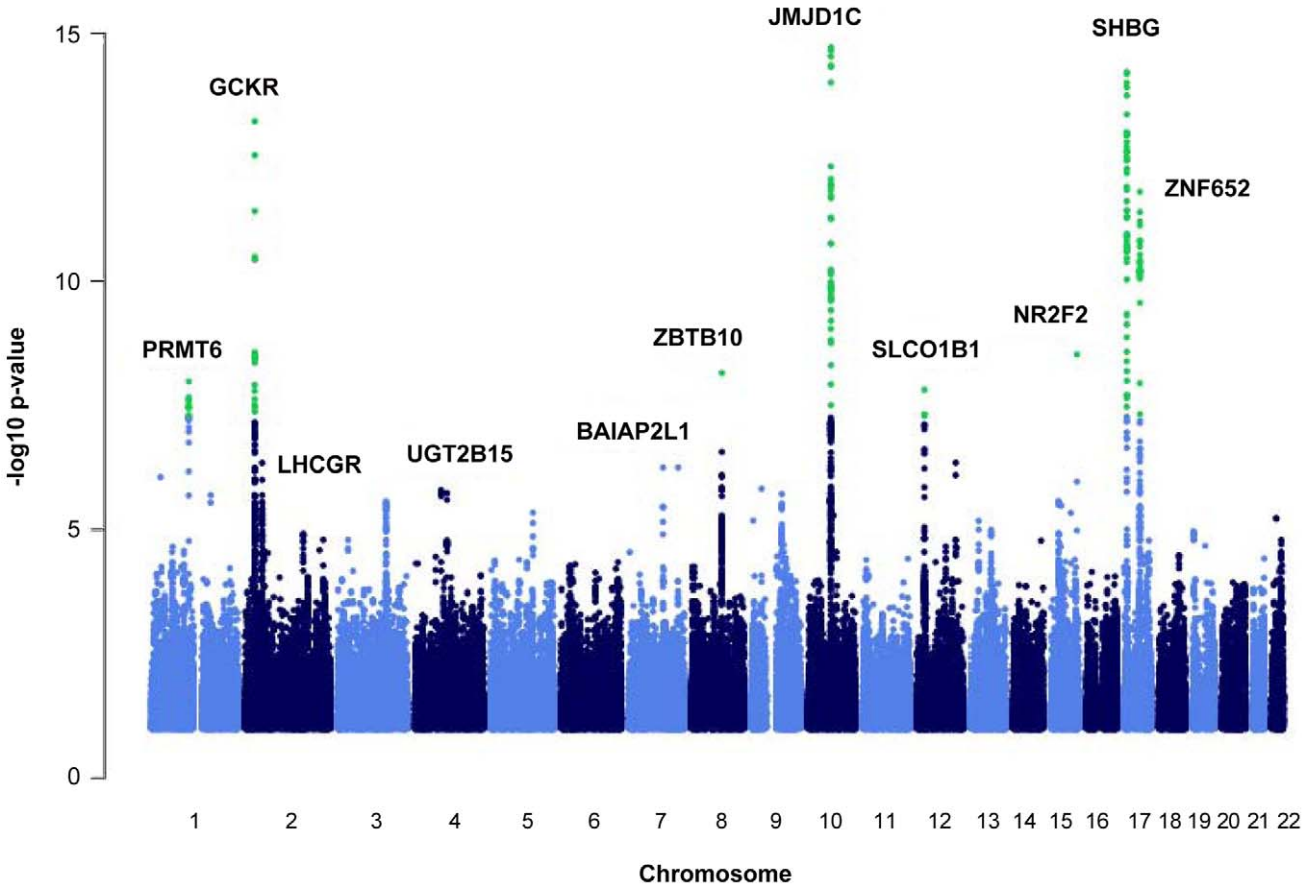


**Figure 1. Manhattan plot of the autosomal SNPs identified in the GWA meta-analysis.** The Manhattan plot depicts the SNPs identified in the GWAS analysis labeled with the nearest gene on the plot. The SNP identified on the X chromosome, rs1573036, at Xq22.3, is not included in this figure.
doi:10.1371/journal.pgen.1002805.g001

component respectively. The *SHBG* locus accounted for ~10% and ~6.6% of the genetic variation in men and women respectively with the lead SNP in isolation accounting for ~7.8% and ~3.3% of the variation in men and women, respectively.

We identified genes near the associated SNPs and explored their biologic relevance to SHBG. The genes associated with identified SNPs included the *SHBG* locus (rs12150660, 17p13.1, $p = 1.8\times10^{-106}$), *PRMT6* (rs17496332, 1p13.3, $p = 1.4\times10^{-11}$), *GCKR* (rs780093, 2p23.3, $p = 2.2\times10^{-16}$), *ZBTB10* (rs440837, 8q21.13, $p = 3.4\times10^{-09}$), *JMJD1C* (rs7910927, 10q21.3, $p = 6.1\times10^{-35}$), *SLCO1B1* (rs4149056, 12p12.1, $p = 1.9\times10^{-08}$), *NR2F2* (rs8023580, 15q26.2, $p = 8.3\times10^{-12}$), *ZNF652* (rs2411984, 17q21.32, $p = 3.5\times10^{-14}$), *TDGF3* (rs1573036, Xq22.3, $p = 4.1\times10^{-14}$), *LHCGR* (rs10454142, 2p16.3, $p = 1.3\times10^{-07}$), *BAIAP2L1* (rs3779195, 7q21.3, $p = 2.7\times10^{-08}$), and *UGT2B15* (rs293428, 4q13.2, $p = 5.5\times10^{-06}$) (Figure 1).

We used the online tool STRING (www.string-db.org) to perform pathway analyses to explore possible interactions between the *SHBG* gene and the proteins encoded by the 11 most plausible genes nearest the 11 SNPs listed above. There was an interaction noted between *GCKR* and *JMJD1C* which were associated with the lipoprotein fractions VLDL and HDL, respectively [11]. In an expanded analysis, we assessed protein interactions among *SHBG* and 67 genes within 500 kb of our 11 identified SNPs and uncovered additional protein interaction pathways. An interaction between two proteins encoded by *GTF2A1L* and *STON1* was found; these proteins are co-expressed in testicular germ cells in the mouse [12]. An interaction between *LHCGR* and *BRI3* encoded proteins that are associated with the G-protein coupled receptor complex in the human luteinizing hormone receptor was also identified [13]. Finally, an interaction between *LHCGR* and *IAPP* (amylin) proteins which are components of a ligand/G-protein receptor/G-protein alpha subunit complex was found (database: www.reactome.com).

Targeted analysis of two strong candidate genes, hepatocyte nuclear factor-4α (*HNF4α*) and peroxisome-proliferating receptor γ (*PPARγ*) did not identify any SNPs at *HNF4α* but did identify one SNP, rs2920502, at *PPARγ* that reached statistical significance ($p = 9.9\times10^{-5}$) and a second SNP at *PPARγ*, rs13081389, that reached nominal significance ($p = 0.01$).

## Discussion

In total, we identified 12 genomic regions associated with circulating SHBG concentrations, including extensive allelic heterogeneity at the *SHBG* locus itself. Conditional meta-analyses carried out at the *SHBG* locus, identified nine genome-wide significant SNPs with low correlation ($r^2<0.01$) between them. Two of these signals (rs6258 [10] and rs6259) are missense variants and two are low frequency variants (MAF ~2%). Furthermore, rs12150660 is highly correlated ($r^2>0.95$) [10] with a pentanu-

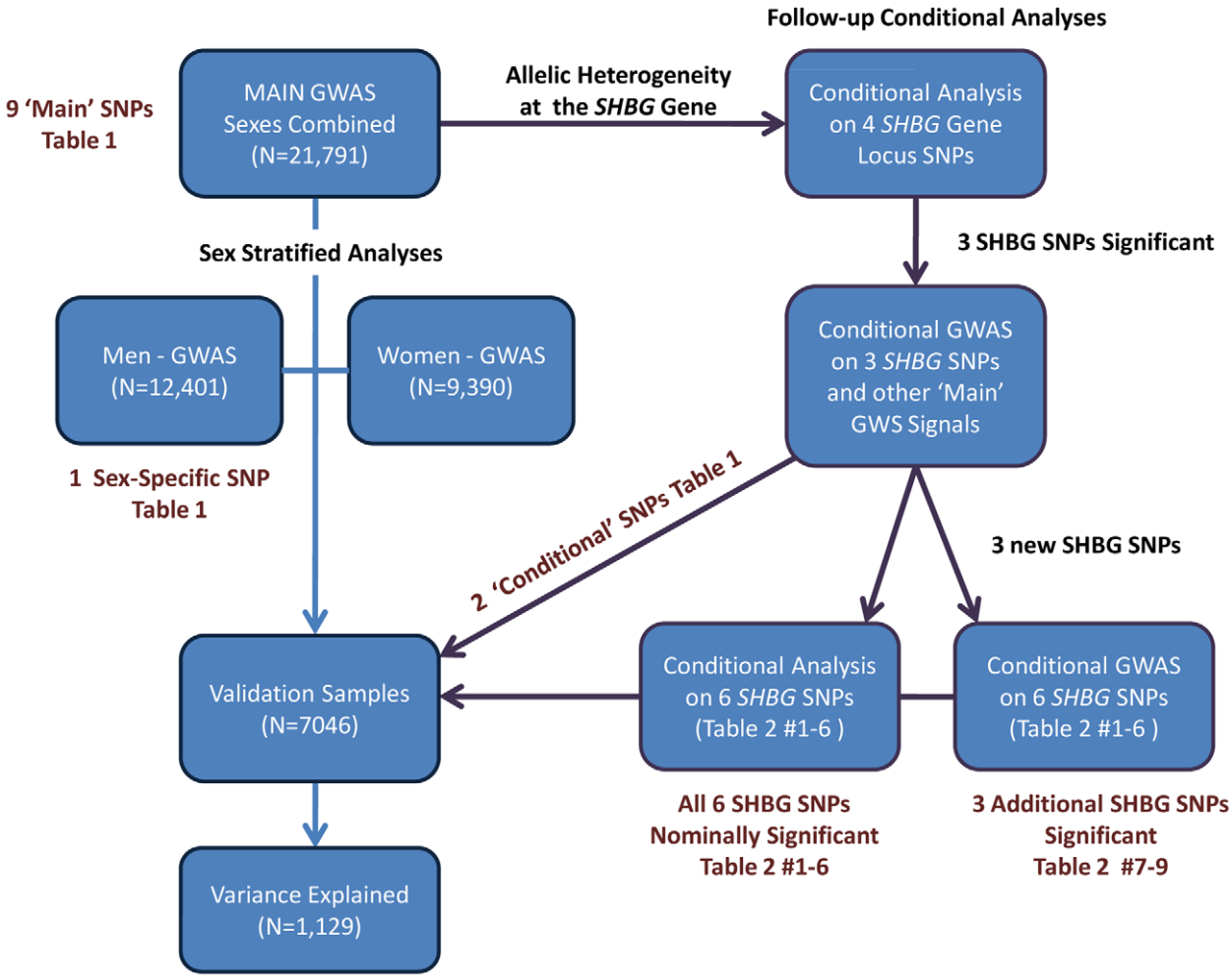


**Figure 2. Summary of the analytic plan.**
doi:10.1371/journal.pgen.1002805.g002

cleotide repeat, which affects *SHBG* expression *in-vitro* [14]. To our knowledge, the magnitude of secondary signals observed at this locus are the largest seen for any complex trait.

The proportion of genetic variance in SHBG serum concentrations explained when accounting for sex specific effects, the multiple signals of association at the *SHBG* locus, and the

**Table 2.** Statistically independent signals at the *SHBG* gene locus.

| SNP # | Model | Conditioned On SNP # | SNP | Position | Effect Allele | Other Allele | EAF | Beta | SE | p-value | Discovery p-value | Discovery Beta |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Full model | 2–6 | rs12150660 | 7462640 | t | g | 0.24 | 0.082 | 0.005 | 1.89E-55 | 1.19E-79 | 0.10 |
| 2 | Full model | 1,3–6 | rs6258 | 7475403 | t | c | 0.02 | −0.272 | 0.017 | 1.03E-60 | 2.69E-46 | −0.2613 |
| 3 | Full model | 1–2,4–6 | rs1641537 | 7486446 | t | c | 0.14 | −0.064 | 0.006 | 1.20E-24 | 8.19E-39 | −0.0814 |
| 4 | Full model | 1–3,5–6 | rs1625895 | 7518840 | t | c | 0.12 | −0.06 | 0.006 | 1.75E-21 | 1.17E-14 | −0.052 |
| 5 | Full model | 1–4,6 | rs6259 | 7477252 | a | g | 0.11 | 0.026 | 0.007 | 0.0001 | 1.46E-07 | 0.0372 |
| 6 | Full model | 1–5 | rs10432029 | 7331393 | a | g | 0.79 | 0.0136 | 0.006 | 0.01 | 7.52E-16 | 0.0446 |
| 7 | Conditional | 1–6 | rs9901675 | 7425536 | a | g | 0.05 | −0.057 | 0.01 | 1.46E-07 | 5.2E-12 | −0.07 |
| 8 | Conditional | 1–6 | rs8077824 | 7588951 | a | g | 0.02 | 0.075 | 0.018 | 4.58E-05 | 0.01 | 0.0451 |
| 9 | Conditional | 1–6 | rs9303218 | 7339386 | t | c | 0.77 | 0.026 | 0.006 | 9,89E-06 | 1.21E-11 | 0.0344 |

All SNPs are on the+strand and positions are based on build 36. EAF = 'effect allele frequency'. Beta units are per-allele effect estimates in natural log transformed nmol/L. 'Full model' SNPs were all included in a single regression model, where the effect estimates for each SNP are adjusted for the effect of the others in the model. 'Conditional' SNPs are SNPs with low pair-wise LD (HapMap $r^2<0.01$) that were identified after conditioning on the full model SNPs.
doi:10.1371/journal.pgen.1002805.t002

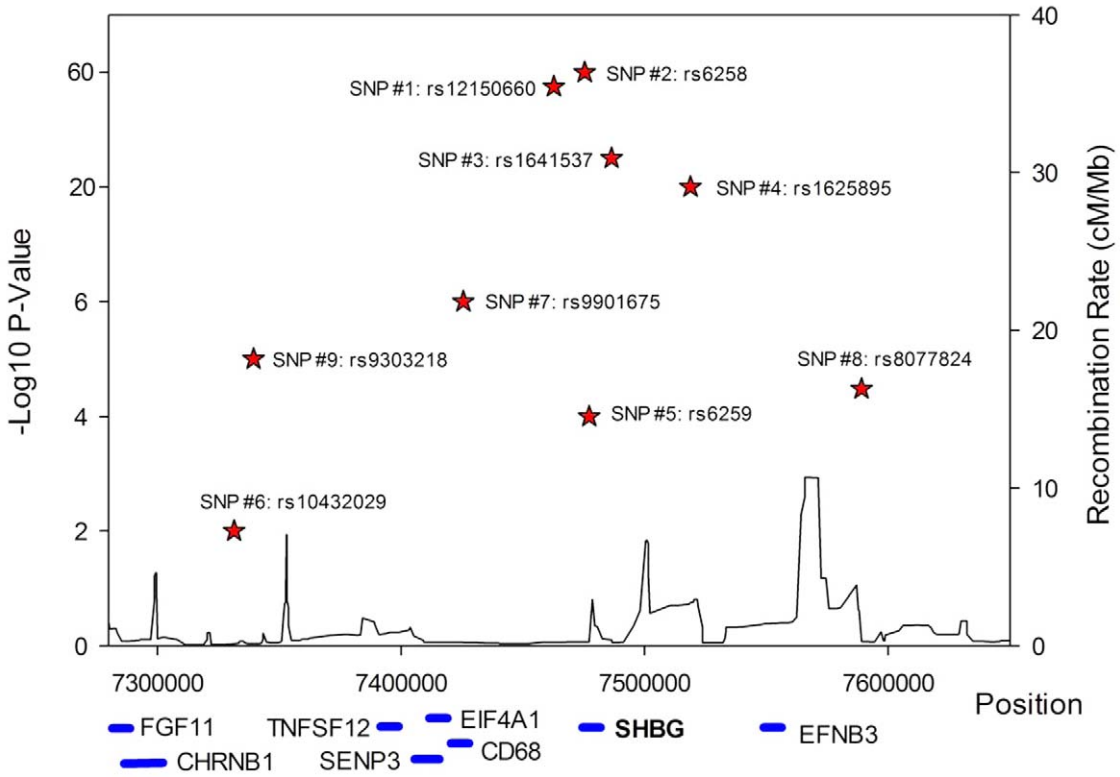


**Figure 3. Allelic heterogeneity at the *SHBG* gene locus.** There was significant allelic heterogeneity at the SHBG gene locus. The nine independent signals identified in the SHBG gene are shown in relation to their position within the gene. All positions based on build 36. Not all genes are shown.
doi:10.1371/journal.pgen.1002805.g003

additional *trans* signals identified post conditional analysis was ~15.6% in men and ~8.4% in women. The *SHBG* locus accounted for ~10% and ~6.6% of the genetic variance in men and women, respectively, with the lead SNP explaining most of the genetic variation at ~7.8% for men and ~3.3% for women. Thus additional signals at the *SHBG* locus identified through conditional analyses approximately doubled the variance of the trait explained. While we provide evidence for multiple variants associated with SHBG concentrations, further studies are needed to pinpoint the causal loci and functional variants. For the 11 regions outside the *SHBG* locus, most have biologically plausible related genes within 300 kb.

## Biology of Plausible Genes near Identified SNPs

Several genes near the identified SNPs regulate sex steroid production and function. The *NR2F2* locus (15q26.2) encodes a nuclear receptor important in testicular Leydig cell function, the primary source of gonadal testosterone production [15], and has been linked to male infertility [16]. *NR2F2* has also been associated with estrogen receptor alpha (ERα) signaling and may influence hormone responsivity in breast cancer [17]. *PRMT6* (1p13.3) also encodes a nuclear receptor regulatory protein that mediates estrogen signaling as a co-activator of the estrogen receptor [18]. *LHCGR* (2p16.3) encodes the luteinizing hormone receptor which was associated with polycystic ovary syndrome (PCOS) in a recent GWAS [19,20]. PCOS is both a reproductive and metabolic disorder characterized by higher testosterone serum concentrations as well as an increased prevalence of obesity, insulin resistance, and T2D in women. Inappropriate secretion of luteinizing hormone leads to increased ovarian production of testosterone. Coincident lower SHBG concentrations contribute to increased bioavailable testosterone concentrations and the expression of both reproductive and metabolic phenotypes in PCOS [21,22,23].

The *SLCO1B1* locus encodes a liver-specific transporter of thyroid hormone as well as estrogens which impact liver production of SHBG [24]. *JMJD1C* (10q21.3), also known as *TRIP 8* (thyroid hormone receptor interactor protein 8 [25]), may impact SHBG concentrations via thyroid hormone effects on liver protein production. Thyroid hormone may alter SHBG production through effects on *HNF4*α which is known to regulate *SHBG* transcription [26,27].

Many of the genes identified are involved in carbohydrate and lipid metabolism and liver function. The *GCKR* locus (2p23.3) encodes a protein that regulates glucokinase activity and has been associated with T2D in several ethnic populations [28,29,30,31]. *GCKR* has been associated with metabolic and inflammatory traits including triglyceride concentrations and other lipid fractions [30,32], fasting plasma glucose [33,34], insulin concentrations, uric acid, c-reactive protein (CRP), and non-alcoholic fatty liver disease which are all characteristic of the metabolic syndrome and T2D [28,35,36,37,38,39,40,41,42]. The *SLCO1B1* locus (12p12.1) codes for a protein, hepatocyte protein anion-transporting polypeptide 1B1, involved in liver metabolism of both endogenous and exogenous compounds [43]. Consistent with *SLCO1B1's* role in liver metabolism, the same SNP (rs4149056) has been associated with circulating bilirubin concentrations in previous GWAS [44]. *BAIAP2L1* (7q21.3) encodes a protein important in cytoskeleton organization [45] that has been associated with the inflammatory marker CRP in patients with arthritis [46]. *BAIAP2L1* is also known as *IRTKS* (insulin receptor tyrosine kinase substrate) which

is involved in insulin receptor signaling [47] and may relate to insulin resistant states including obesity and T2D [48,49, 50,51,52,53,54]. We conducted a targeted analysis of *PPARγ*, a gene that influences *SHBG* gene expression in the liver [1,55] and is associated with T2D [56,57]. Our analysis identified one significant SNP (rs2920502, $p = 9.9 \times 10^{-5}$) and a second nominally significant SNP (rs13081389, $p = 0.01$) at *PPARγ*. Some of the identified genes involved in hepatic metabolism of lipids and carbohydrates may be affect SHBG concentrations indirectly through effects on the SHBG transcription regulator *HNF4α* although *HNF4α* itself was not identified in this meta-analyses [27,58,59,60].

The *UGT2B15* locus (4q13.2) was significantly associated with SHBG concentrations in men but not women in this meta-analysis. *UGT2B15* belongs to a family of genes (the UGT2B gene family) that code for enzymes involved in the metabolism of sex hormones through glucuronidation which allows for excretion of sex steroids through the kidney and the gut via bile excretion [61,62], primary clearance mechanisms for sex steroids [63]. UGT2B15 is involved in the conjugation and inactivation of testosterone [64]. An association between rs293428 in the *UGT2B15* locus and circulating SHBG concentrations in men is supported by a previous study demonstrating that a non-synonymous SNP in *UGT2B15* (rs1902023; D85Y) is associated with serum SHBG concentrations in younger adult men [65]. *UGT2B15* is thought to play a significant role in local tissue inactivation of androgens in androgen dependent prostate cancer [66,67]. The mechanism behind the influence of genetic variants in *UGT2B15* on SHBG concentrations is unknown, but one may speculate that *UGT2B15* affects the local androgenic environment in selected tissues, which in turn results in regulation of SHBG concentrations.

In addition to *UGT2B15*, three other genes near the identified SNPs are associated with carcinogenesis, particularly in the prostate and breast. *ZBTB10* (8q21.13), has been linked to breast cancer [68]. In breast cancer cell lines *ZBTB10* is suppressed by ROS-microRNA27a thereby enhancing ERα alpha expression and mediating estrogen effects [17]. The *ZNF652* (17q21.32) locus codes for a DNA binding protein thought to act as a tumor suppressor gene in breast cancer [69,70,71] that is also co-expressed with the androgen receptor in prostate cancer [72]. *TDGF3*, teratocarcinoma derived growth factor 3, is the only significant region identified on the X chromosome ((Xq22.3). *TDGF3* is a *pseudogene* of *TDGF1* located on chromosome 3p23-p21 that has been associated with testicular germ cell tumors [73].

### Strengths and Limitations

This GWAS meta-analysis incorporated data from approximately 22,000 men and women from 16 epidemiologic cohorts. The overall size of the study yields power but the meta-analysis of data from different epidemiologic studies requires the inclusion of different laboratory methods. The different studies used a variety of assay methodologies to measure serum SHBG concentrations although the vast majority were immunoassays (Tables S1 and S2, Text S1) with similar methodologies. Variation introduced by the use of different SHBG assays would result in loss of statistical power and likely bias toward the null. Additionally, the majority of women were post-menopausal as ascertained by self-report in all studies (Table S1). SHBG concentrations, like testosterone, decline only slightly across the menopause [74] so adjustment for menopause status is not necessary. SHBG may also increase with ovulation and be slightly higher in the luteal versus the follicular phase of the menstrual cycle in premenopausal women, but most studies did not collect data on menstrual phase at the time of SHBG measurement so adjustment for menstrual phase was not possible [75]. Finally, individuals were not excluded based on health status, therefore some individuals with chronic conditions that may affect hepatic production of or clearance of proteins including SHBG such as liver disease, renal disease, or severe malnutrition, may have been included in this analysis.

### Conclusion

SHBG synthesis in the liver is known to be affected directly or indirectly by estrogens, androgens and thyroid hormones and has been observed to be inversely associated with the higher insulin concentrations characteristic of insulin resistant states such as T2D [1,6]. In summary, the results of this GWAS reflect these influences. Three regions map to proteins related to hepatic function (12p12.1-*SLCO1B1* [*76*], 2p23.3-*GCKR* [*77*] and 10q21.3-*JMJD1C* [*77*]). In addition, 2p23.3-*GCKR* and 7q21.3-*BAIAP2L1* [alias insulin receptor tyrosine kinase substrate (*IRTKS*)] are involved in susceptibility to T2D [48] and insulin signaling [47], respectively. Two signals also mapped to loci involved in thyroid hormone regulation (10q21.3-*JMJD1C* and 12p12.1-*SLCO1B1*). One signal mapped to the receptor for luteinizing hormone 2p16.3-*LHCGR* [*20*], the hormone that stimulates testosterone production. Five regions mapped to genes previously implicated in androgen and estrogen signaling (1p13.3-*PRMT6* [*18*], 8q21.13-*ZBTB10* [*17*], 12p12.1-*SLCO1B1* [*76*], 15q26.2-*NR2F2* [*78*], 4q13.2-*UGT2B15* [*63*]).

We have combined a conventional GWAS approach with detailed additional analyses, including sex stratification, conditional analysis and imputation from 1000 Genomes. Our results demonstrate that these approaches can lead to an appreciable gain in heritable variance explained. It does however highlight the complexity of elucidating individual variant causality through statistical approaches. In addition to the extensive allelic heterogeneity at the *SHBG* locus, our data identify loci with a role in sex steroid hormone metabolism, which may help elucidate the role of sex steroid hormones in disease, particularly T2D and hormone-sensitive cancers.

## Methods

We performed a genome wide association study (GWAS) meta-analysis of 21,791 individuals (Table S1: 9,390 women, 12,401 men) from ten observational studies. Data from an additional six studies totaling 7,046 individuals (Table S2: 4,509 women; 2,537 men) were used for validation. The proportion of variance explained was estimated in an independent study (InCHIANTI, $n = 1,129$). The individual study protocols were approved by their respective institution's ethics committee/institutional review board and all participants provided informed consent prior to participation. Individuals known to be taking hormonal contraceptives or hormone replacement therapy at time of SHBG measurement were excluded from analysis. Age, sex and body mass index (BMI) were included as covariates. After applying standard quality control measures, imputed genotypes were available for approximately 2.5 M SNPs. See Figure 2 for an overview of the analytic plan and the Text S1 for further information for individual studies included in this meta-analysis.

### GWAS Conditional Meta-Analysis Steps

**Conditional analysis #1.** The initial starting point for the conditional analysis was the four *SHBG* locus SNPs that all showed low Hapmap LD ($r^2 < 0.05$) with each other: rs12150660 (lead SNP Table 1), rs6258 $p = 2.7 \times 10^{-46}$, rs1625895 $p = 1.2 \times 10^{-14}$ and rs3853894 $p = 2.5 \times 10^{-11}$. Each cohort fitted a single

regression model, fitting SHBG concentrations against these four genome-wide significant SHBG locus SNPs (rs12150660, rs6258, rs1625895 and rs3853894), in addition to age, sex and BMI. After meta-analyzing the results from all cohorts, three of the SNPs retained genome wide significance when regressed against each other, with the fourth SNP narrowly missing that threshold (rs3853894, $p = 4.1 \times 10^{-6}$).

**Conditional GWAS #1 (Table 1, conditional analysis).** We next performed a conditional GWAS meta-analysis, where each study included, as additional covariates to the original analysis plan, the ten genome-wide significant autosomal SNPs (the eight 'Main' signals from Table 1 and the two unique *SHBG* locus signals described above in addition to the lead SNP rs12150660: rs6258 and rs1625895). Three additional signals (independence based on HapMap $r^2<0.05$) at the *SHBG* locus reached genome-wide significance (rs1641537 $p = 7.8 \times 10^{-32}$, rs6259 $p = 1.5 \times 10^{-12}$ and rs10432029 $p = 3 \times 10^{-8}$), giving a total of six independent signals in this gene region. In addition, two novel signals reached genome-wide significance in the conditional analysis, at 7q21.3 (rs3779195 $p = 1 \times 10^{-8}$) and 2p16.3 (rs10454142 $p = 3 \times 10^{-8}$). After replication, only rs3779195 at the *BAIAP2L1* locus retained genome-wide significance.

**Conditional analysis #2 (Table 2, full model).** Given the six signals observed at the *SHBG* locus (three through conditional analysis #1 rs12150660, rs6258, rs1625895, three through LD estimates from conditional GWAS #1: rs1641537, rs6259, rs10432029), we sought to confirm which of these six were truly independent by a second round of conditional analysis. All discovery and replication cohorts fitted a single regression model of the six SNPs (SNPs # 1–6, Table 2) against SHBG concentrations, using the same parameters and covariates as conditional analysis #1. Four of the six SNPs (#1–4: rs12150660, rs6258, rs1641537, and rs1625895) retained genome-wide significance when conditioned against each other, with two showing nominal evidence of association (SNP #5 rs6259, $p = 0.0001$; SNP #6 rs10432029, $p = 0.01$).

**Conditional GWAS #2 (Table 2, conditional model).** Finally, we performed a second conditional GWAS analysis, adjusting for the six *SHBG* locus SNPs which had evidence of association from conditional analysis #2. All the discovery cohorts were used in this analysis, in addition to three replication cohorts (total sample size 24,354). This analysis revealed evidence for a further three independent signals at the *SHBG* locus (based on HapMap $r^2<0.01$), SNP #7 rs9901675 $p = 1.5 \times 10^{-7}$, SNP #8 rs8077824 $p = 4.6 \times 10^{-5}$, and SNP #9 rs9393218 $p = 9.9 \times 10^{-6}$.

### Sensitivity Analysis—Allelic Heterogeneity at the SHBG Locus

We performed a sensitivity analysis using samples from the 1966 Northern Finland Birth Cohort (NFBC1966) study to further investigate allelic heterogeneity at the *SHBG* locus (Text S1). The conditional meta-analysis showed evidence for up to nine signals at the *SHBG* locus, but it is possible that these signals could be explaining a much smaller number of causal variants in the region. Since 1000 Genomes imputation allows us to assess the genetic variation associated with a phenotype across a much denser set of markers, it increases our power to detect allelic heterogeneity within a region. Therefore, 1000 Genomes imputation was carried out on all the samples in the NFBC1966 study and forward selection was used to identify the set of SNPs that best explain the variation in the SHBG phenotype. 1000 Genomes imputation was carried out using IMPUTE2. The mean genotype probabilities for each SNP were calculated and used in the model selection step. Only SNPs 250 kb upstream and 250 kb downstream from the *SHBG* locus (7283453–7786700 bp) were used in the analysis. All SNPs with MAF <0.1% or an imputation quality score less than 0.4 were excluded from the analysis. In total, 1978 *SHBG* region SNPs measured or imputed in 4467 samples from the NFBC1966 study were used in the sensitivity analysis. Forward selection was implemented in R (version 2.13.0) using the stepAIC package to estimate the Akaikie Information Criterion (AIC), an inclusion parameter. Given the high degree of correlation between the SNPs in this region, we increased the penalty (k) on the number of terms included in the model to 12 (where it is usually two), to minimize possible over fitting. The final model included seven SNPs, adjusted for sex and BMI.

### Pathway Analysis

We examined potential interactions among the proteins encoded by the *SHBG* locus and the proteins encoded by the 11 genes (*ZBT10*, *TDGF1*, *ZNF652*, *PRMT6*, *JMJD1C*, *GCKR*, *BAIAP2L1*, *LHCGR*, *SLCO1B1*, *UGT2B15*, *NR2F2*) closest to the 11 identified SNPs using pathway analysis with Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) Pathways Analysis (www.string-db.org). The interactions explored by STRING include direct (physical) and indirect (functional) associations. We then expanded the analysis to examine protein interactions among the *SHBG* gene and the proteins encoded by 67 genes within 500 kb of the 11 identified SNPs.

### Targeted Candidate Gene Analysis

We conducted targeted analysis of two strong candidate genes, hepatocyte nuclear factor-4α (*HNF4α*) and peroxisome-proliferating receptor γ (*PPARγ*). Statistical significance thresholds were set correcting for the number of SNPs tested in each gene region (±100 kb).

## Supporting Information

**Table S1** Characteristics of 21,791 individuals from 10 discovery cohorts included in the meta-analysis.
(DOC)

**Table S2** Characteristics of 8,175 individuals from the six cohorts included in the validation analysis (WHI, CARDIA, Prospect-EPIC, MrOs, NHS, YFS) and the independent cohort used to estimate the proportion of genetic variance explained by the indentified SNPs (InChianti).
(DOC)

**Table S3** Hapmap (release 22) linkage disequilibrium estimates for the nine *SHBG* gene locus single nucleotide polymorphisms.
(DOC)

**Text S1** Supplementary Methods with Specific Cohort Information.
(DOC)

## Acknowledgments

We are grateful to all study participants and staff in our participating studies. Full study acknowledgements are available in Text S1.

## Author Contributions

Conceived and designed the experiments: FH de Jong, O Raitakari, A Teumer, C Ohlsson, JM Murabito, JRB Perry. Analyzed the data: JRB Perry, AD Coviello, R Haring, M Wellons, D Vaidya, T Lehtimäki, S Keildson, KL Lunetta, C He, M-R Järvelin. Wrote the paper: AD Coviello, R Haring, M Wellons, D Vaidya, T Lehtimäki, S Keildson, KL Lunetta, C He, TM Frayling, A Murray, S Franks, M-R Järvelin, FH de

Jong, O Raitakari, A Teumer, C Ohlsson, JM Murabio, JRB Perry. Individual cohort phenotyping/genotyping: KL Lunetta, C He, M Fornage, V Lagou, M Mangino, NC Onland-Moret, B Chen, J Eriksson, M Garcia, YM Liu, A Koster, K Lohman, L-P Lyytikäinen, A-K Petersen, J Prescott, L Stolk, L Vandenput, AR Wood, WV Zhuang, A Ruokonen, A-L Hartikainen, A Pouta, S Bandinelli, AD Coviello, R Biffar, G Brabant, DG Cox, Y Chen, S Cummings, L Ferrucci, MJ Gunter, SE Hankinson, H Martikainen, A Hofman, G Homuth, T Illig, J-O Jansson, AD Johnson, D Karasik, M Karlsson, J Kettunen, DP Kiel, P Kraft, J Liu, Ö Ljunggren, M Lorentzon, M Maggio, MRP Markus, D Mellström, I Miljkovic, D Mirel, S Nelson, L Morin Papunen, PHM Peeters, I Prokopenko, L Raffel, M Reincke, AP Reiner, K Rexrode, F Rivadeneira, SM Schwartz, D Siscovick, N Soranzo, D Stöckl, S Tworoger, AG Uitterlinden, CH van Gils, RS Vasan, H-E Wichmann, G Zhai, S Bhasin, M Bidlingmaier, SJ Chanock, I De Vivo, TB Harris, DJ Hunter, M Kähönen, S Liu, P Ouyang, TD Spector, YT van der Schouw, J Viikari, H Wallaschofski, MI McCarthy, C Ohlsson, TM Frayling, A Murray, S Franks, M-R Järvelin.

## References


1. Hammond GL (2011) Diverse roles for sex hormone-binding globulin in reproduction. Biol Reprod 85: 431–441.
2. Hammes A, Andreassen TK, Spoelgen R, Raila J, Hubner N, et al. (2005) Role of endocytosis in cellular uptake of sex steroids. Cell 122: 751–762.
3. Thompson DJ, Healey CS, Baynes C, Kalmyrzaev B, Ahmed S, et al. (2008) Identification of common variants in the SHBG gene affecting sex hormone-binding globulin levels and breast cancer risk in postmenopausal women. Cancer Epidemiol Biomarkers Prev 17: 3490–3498.
4. Wassell J, Michail M, Soliman N, Wardle PG (2011) The value of sex hormone binding globulin (SHBG) in predicting treatment response in polycystic ovary syndrome (PCOS). Clin Lab 57: 95–98.
5. Xita N, Tsatsoulis A, Chatzikyriakidou A, Georgiou I (2003) Association of the (TAAAA)n repeat polymorphism in the sex hormone-binding globulin (SHBG) gene with polycystic ovary syndrome and relation to SHBG serum levels. J Clin Endocrinol Metab 88: 5976–5980.
6. Ding EL, Song Y, Manson JE, Hunter DJ, Lee CC, et al. (2009) Sex hormone-binding globulin and risk of type 2 diabetes in women and men. N Engl J Med 361: 1152–1163.
7. Perry JR, Weedon MN, Langenberg C, Jackson AU, Lyssenko V, et al. (2010) Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. Hum Mol Genet 19: 535–544.
8. Coviello AD, Zhuang WV, Lunetta KL, Bhasin S, Ulloor J, et al. (2011) Circulating testosterone and SHBG concentrations are heritable in women: the Framingham Heart Study. J Clin Endocrinol Metab 96: E1491–1495.
9. Melzer D, Perry JR, Hernandez D, Corsi AM, Stevens K, et al. (2008) A genome-wide association study identifies protein quantitative trait loci (pQTLs). PLoS Genet 4: e1000072. doi:10.1371/journal.pgen.1000072
10. Ohlsson C, Wallaschofski H, Lunetta KL, Stolk L, Perry JR, et al. (2011) Genetic determinants of serum testosterone concentrations in men. PLoS Genet 7: e1002313. doi:10.1371/journal.pgen.1002313
11. Chasman DI, Pare G, Mora S, Hopewell JC, Peloso G, et al. (2009) Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. PLoS Genet 5: e1000730. doi:10.1371/journal.pgen.1000730
12. Han SY, Zhou L, Upadhyaya A, Lee SH, Parker KL, et al. (2001) TFIIAalpha/beta-like factor is encoded by a germ cell-specific gene whose expression is up-regulated with other general transcription factors during spermatogenesis in the mouse. Biol Reprod 64: 507–517.
13. Kudo M, Osuga Y, Kobilka BK, Hsueh AJ (1996) Transmembrane regions V and VI of the human luteinizing hormone receptor are required for constitutive activation by a mutation in the third intracellular loop. J Biol Chem 271: 22470–22478.
14. Hogeveen KN, Talikka M, Hammond GL (2001) Human sex hormone-binding globulin promoter activity is influenced by a (TAAAA)n repeat element within an Alu sequence. J Biol Chem 276: 36383–36390.
15. Martin LJ, Tremblay JJ (2010) Nuclear receptors in Leydig cell gene expression and function. Biol Reprod 83: 3–14.
16. Hu Z, Xia Y, Guo X, Dai J, Li H, et al. (2011) A genome-wide association study in Chinese men identifies three risk loci for non-obstructive azoospermia. Nat Genet 44: 183–186.
17. Li X, Mertens-Talcott SU, Zhang S, Kim K, Ball J, et al. (2010) MicroRNA-27a Indirectly Regulates Estrogen Receptor {alpha} Expression and Hormone Responsiveness in MCF-7 Breast Cancer Cells. Endocrinology 151: 2462–2473.
18. Harrison MJ, Tang YH, Dowhan DH (2010) Protein arginine methyltransferase 6 regulates multiple aspects of gene expression. Nucleic Acids Res 38: 2201–2216.
19. Franks S, Gharani N, McCarthy M (2001) Candidate genes in polycystic ovary syndrome. Hum Reprod Update 7: 405–410.
20. Chen ZJ, Zhao H, He L, Shi Y, Qin Y, et al. (2011) Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. Nat Genet 43: 55–59.
21. Cousin P, Calemard-Michel L, Lejeune H, Raverot G, Yessaad N, et al. (2004) Influence of *SHBG* gene pentanucleotide TAAAA repeat and D327N polymorphism on serum sex hormone-binding globulin concentration in hirsute women. J Clin Endocrinol Metab 89: 917–924.
22. Ehrmann DA (2005) Polycystic ovary syndrome. N Engl J Med 352: 1223–1236.
23. Franks S (1995) Polycystic ovary syndrome. N Engl J Med 333: 853–861.
24. Abe T, Kakyo M, Tokui T, Nakagomi R, Nishio T, et al. (1999) Identification of a novel gene family encoding human liver-specific organic anion transporter LST-1. J Biol Chem 274: 17159–17163.
25. Lee JW, Choi HS, Gyuris J, Brent R, Moore DD (1995) Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor. Mol Endocrinol 9: 243–254.
26. Selva DM, Hammond GL (2009) Thyroid hormones act indirectly to increase sex hormone-binding globulin production by liver via hepatocyte nuclear factor-4alpha. J Mol Endocrinol 43: 19–27.
27. Janne M, Hammond GL (1998) Hepatocyte nuclear factor-4 controls transcription from a TATA-less human sex hormone-binding globulin gene promoter. J Biol Chem 273: 34105–34114.
28. Reiling E, van 't Riet E, Groenewoud MJ, Welschen LM, van Hove EC, et al. (2009) Combined effects of single-nucleotide polymorphisms in *GCK*, *GCKR*, *G6PC2* and *MTNR1B* on fasting plasma glucose and type 2 diabetes risk. Diabetologia 52: 1866–1870.
29. Qi Q, Wu Y, Li H, Loos RJ, Hu FB, et al. (2009) Association of *GCKR* rs780094, alone or in combination with *GCK* rs1799884, with type 2 diabetes and related traits in a Han Chinese population. Diabetologia 52: 834–843.
30. Ling Y, Li X, Gu Q, Chen H, Lu D, et al. (2011) Associations of common polymorphisms in *GCKR* with type 2 diabetes and related traits in a Han Chinese population: a case-control study. BMC Med Genet 12: 66.
31. Tanaka D, Nagashima K, Sasaki M, Yamada C, Funakoshi S, et al. (2011) *GCKR* mutations in Japanese families with clustered type 2 diabetes. Mol Genet Metab 102: 453–460.
32. Varbo A, Benn M, Tybjaerg-Hansen A, Grande P, Nordestgaard BG (2011) *TRIB1* and *GCKR* polymorphisms, lipid levels, and risk of ischemic heart disease in the general population. Arterioscler Thromb Vasc Biol 31: 451–457.
33. Takeuchi F, Katsuya T, Chakrewarthy S, Yamamoto K, Fujioka A, et al. (2010) Common variants at the *GCK*, *GCKR*, *G6PC2-ABCB11* and *MTNR1B* loci are associated with fasting glucose in two Asian populations. Diabetologia 53: 299–308.
34. Onuma H, Tabara Y, Kawamoto R, Shimizu I, Kawamura R, et al. (2010) The GCKR rs780094 polymorphism is associated with susceptibility of type 2 diabetes, reduced fasting plasma glucose levels, increased triglycerides levels and lower HOMA-IR in Japanese population. J Hum Genet 55: 600–604.
35. Hadarits F, Kisfali P, Mohas M, Maasz A, Duga B, et al. (2011) Common functional variants of *APOA5* and *GCKR* accumulate gradually in association with triglyceride increase in metabolic syndrome patients. Mol Biol Rep.
36. Hu C, Zhang R, Wang C, Yu W, Lu J, et al. (2010) Effects of *GCK*, *GCKR*, *G6PC2* and *MTNR1B* variants on glucose metabolism and insulin secretion. PLoS ONE 5: e11761. doi:10.1371/journal.pone.0011761
37. Kozian DH, Barthel A, Cousin E, Brunnhofer R, Anderka O, et al. (2010) Glucokinase-activating GCKR polymorphisms increase plasma levels of triglycerides and free fatty acids, but do not elevate cardiovascular risk in the Ludwigshafen Risk and Cardiovascular Health Study. Horm Metab Res 42: 502–506.
38. Mohas M, Kisfali P, Jaromi L, Maasz A, Feher E, et al. (2010) GCKR gene functional variants in type 2 diabetes and metabolic syndrome: do the rare variants associate with increased carotid intima-media thickness? Cardiovasc Diabetol 9: 79.
39. Orho-Melander M, Melander O, Guiducci C, Perez-Martinez P, Corella D, et al. (2008) Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and C-reactive protein but lower fasting glucose concentrations. Diabetes 57: 3112–3121.
40. Perez-Martinez P, Delgado-Lista J, Garcia-Rios A, Mc Monagle J, Gulseth HL, et al. (2011) Glucokinase regulatory protein genetic variant interacts with omega-3 PUFA to influence insulin resistance and inflammation in metabolic syndrome. PLoS ONE 6: e20555. doi:10.1371/journal.pone.0020555
41. Tam CH, Ma RC, So WY, Wang Y, Lam VK, et al. (2009) Interaction effect of genetic polymorphisms in glucokinase (*GCK*) and glucokinase regulatory protein (*GCKR*) on metabolic traits in healthy Chinese adults and adolescents. Diabetes 58: 765–769.
42. Yang Z, Wen J, Tao X, Lu B, Du Y, et al. (2011) Genetic variation in the GCKR gene is associated with non-alcoholic fatty liver disease in Chinese people. Mol Biol Rep 38: 1145–1150.
43. Seithel A, Klein K, Zanger UM, Fromm MF, Konig J (2008) Non-synonymous polymorphisms in the human *SLCO1B1* gene: an in vitro analysis of SNP c.1929A>C. Mol Genet Genomics 279: 149–157.

44. Buch S, Schafmayer C, Volzke H, Seeger M, Miquel JF, et al. (2010) Loci from a genome-wide analysis of bilirubin levels are associated with gallstone risk and composition. Gastroenterology 139: 1942–1951 e1942.
45. Miyahara A, Okamura-Oho Y, Miyashita T, Hoshika A, Yamada M (2003) Genomic structure and alternative splicing of the insulin receptor tyrosine kinase substrate of 53-kDa protein. J Hum Genet 48: 410–414.
46. Galligan CL, Baig E, Bykerk V, Keystone EC, Fish EN (2007) Distinctive gene expression signatures in rheumatoid arthritis synovial tissue fibroblast cells: correlates with disease activity. Genes Immun 8: 480–491.
47. Millard TH, Dawson J, Machesky LM (2007) Characterisation of IRTKS, a novel IRSp53/MIM family actin regulator with distinct filament bundling properties. J Cell Sci 120: 1663–1672.
48. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, et al. (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 42: 105–116.
49. Wang L, Muromoto N, Hayashi H, Mitani Y, Uehara H, et al. (1997) Hyperinsulinemia but no diabetes in transgenic mice homozygously expressing the tyrosine kinase-deficient human insulin receptor. Biochem Biophys Res Commun 240: 446–451.
50. Hotamisligil GS, Budavari A, Murray D, Spiegelman BM (1994) Reduced tyrosine kinase activity of the insulin receptor in obesity-diabetes. Central role of tumor necrosis factor-alpha. J Clin Invest 94: 1543–1549.
51. Kusari J, Takata Y, Hatada E, Freidenberg G, Kolterman O, et al. (1991) Insulin resistance and diabetes due to different mutations in the tyrosine kinase domain of both insulin receptor gene alleles. J Biol Chem 266: 5260–5267.
52. Block NE, Komori K, Robinson KA, Dutton SL, Lam CF, et al. (1991) Diabetes-associated impairment of hepatic insulin receptor tyrosine kinase activity: a study of mechanisms. Endocrinology 128: 312–322.
53. Odawara M, Kadowaki T, Yamamoto R, Shibasaki Y, Tobe K, et al. (1989) Human diabetes associated with a mutation in the tyrosine kinase domain of the insulin receptor. Science 245: 66–68.
54. Arner P, Pollare T, Lithell H, Livingston JN (1987) Defective insulin receptor tyrosine kinase in human skeletal muscle in obesity and type 2 (non-insulin-dependent) diabetes mellitus. Diabetologia 30: 437–440.
55. Selva DM, Hammond GL (2009) Peroxisome-proliferator receptor gamma represses hepatic sex hormone-binding globulin expression. Endocrinology 150: 2183–2189.
56. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, et al. (2000) The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 26: 76–80.
57. Hakonarson H, Grant SF (2011) GWAS and its impact on elucidating the etiology of diabetes. Diabetes Metab Res Rev.
58. Selva DM, Hogeveen KN, Innis SM, Hammond GL (2007) Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. J Clin Invest 117: 3979–3987.
59. Shih DQ, Bussen M, Sehayek E, Ananthanarayanan M, Shneider BL, et al. (2001) Hepatocyte nuclear factor-1alpha is an essential regulator of bile acid and plasma cholesterol metabolism. Nat Genet 27: 375–382.
60. Yin L, Ma H, Ge X, Edwards PA, Zhang Y (2011) Hepatic hepatocyte nuclear factor 4alpha is essential for maintaining triglyceride and cholesterol homeostasis. Arterioscler Thromb Vasc Biol 31: 328–336.
61. Wilson W, 3rd, Pardo-Manuel de Villena F, Lyn-Cook BD, Chatterjee PK, Bell TA, et al. (2004) Characterization of a common deletion polymorphism of the *UGT2B17* gene linked to *UGT2B15*. Genomics 84: 707–714.
62. Turgeon D, Carrier JS, Levesque E, Hum DW, Belanger A (2001) Relative enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members. Endocrinology 142: 778–787.
63. Menard V, Eap O, Harvey M, Guillemette C, Levesque E (2009) Copy-number variations (CNVs) of the human sex steroid metabolizing genes *UGT2B17* and *UGT2B28* and their associations with a *UGT2B15* functional polymorphism. Hum Mutat 30: 1310–1319.
64. Levesque E, Beaulieu M, Green MD, Tephly TR, Belanger A, et al. (1997) Isolation and characterization of UGT2B15(Y85): a UDP-glucuronosyltransferase encoded by a polymorphic gene. Pharmacogenetics 7: 317–325.
65. Swanson C, Mellstrom D, Lorentzon M, Vandenput L, Jakobsson J, et al. (2007) The uridine diphosphate glucuronosyltransferase 2B15 D85Y and 2B17 deletion polymorphisms predict the glucuronidation pattern of androgens and fat mass in men. J Clin Endocrinol Metab 92: 4878–4882.
66. Chouinard S, Barbier O, Belanger A (2007) UDP-glucuronosyltransferase 2B15 (UGT2B15) and UGT2B17 enzymes are major determinants of the androgen response in prostate cancer LNCaP cells. J Biol Chem 282: 33466–33474.
67. Chouinard S, Pelletier G, Belanger A, Barbier O (2004) Cellular specific expression of the androgen-conjugating enzymes UGT2B15 and UGT2B17 in the human prostate epithelium. Endocr Res 30: 717–725.
68. Pathi SS, Jutooru I, Chadalapaka G, Sreevalsan S, Anand S, et al. (2011) GT-094, a NO-NSAID, inhibits colon cancer cell growth by activation of a reactive oxygen species-microRNA-27a: ZBTB10-specificity protein pathway. Mol Cancer Res 9: 195–202.
69. Kumar R, Manning J, Spendlove HE, Kremmidiotis G, McKirdy R, et al. (2006) ZNF652, a novel zinc finger protein, interacts with the putative breast tumor suppressor CBFA2T3 to repress transcription. Mol Cancer Res 4: 655–665.
70. Kumar R, Selth L, Schulz RB, Tay BS, Neilsen PM, et al. (2011) Genome-wide mapping of ZNF652 promoter binding sites in breast cancer cells. J Cell Biochem.
71. Kumar R, Cheney KM, Neilsen PM, Schulz RB, Callen DF (2010) CBFA2T3-ZNF651, like CBFA2T3-ZNF652, functions as a transcriptional corepressor complex. FEBS Lett 584: 859–864.
72. Callen DF, Ricciardelli C, Butler M, Stapleton A, Stahl J, et al. (2010) Co-expression of the androgen receptor and the transcription factor ZNF652 is related to prostate cancer outcome. Oncol Rep 23: 1045–1052.
73. Baldassarre G, Romano A, Armenante F, Rambaldi M, Paoletti I, et al. (1997) Expression of teratocarcinoma-derived growth factor-1 (*TDGF-1*) in testis germ cell tumors and its effects on growth and differentiation of embryonal carcinoma cell line NTERA2/D1. Oncogene 15: 927–936.
74. Burger HG, Dudley EC, Cui J, Dennerstein L, Hopper JL (2000) A prospective longitudinal study of serum testosterone, dehydroepiandrosterone sulfate, and sex hormone-binding globulin levels through the menopause transition. J Clin Endocrinol Metab 85: 2832–2838.
75. Schijf CP, van der Mooren MJ, Doesburg WH, Thomas CM, Rolland R (1993) Differences in serum lipids, lipoproteins, sex hormone binding globulin and testosterone between the follicular and the luteal phase of the menstrual cycle. Acta Endocrinol (Copenh) 129: 130–133.
76. Link E, Parish S, Armitage J, Bowman L, Heath S, et al. (2008) *SLCO1B1* variants and statin-induced myopathy–a genomewide study. N Engl J Med 359: 789–799.
77. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, et al. (2011) Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 43: 1131–1138.
78. Lee DK, Kurihara I, Jeong JW, Lydon JP, DeMayo FJ, et al. (2010) Suppression of ERalpha activity by COUP-TFII is essential for successful implantation and decidualization. Mol Endocrinol 24: 930–940.